(12) UK Patent Application (19) GB (11) 2 256 351 (13) A

(43) Date of A publication 02.12.1992

(21) Application No 9111380.3

(22) Date of filing 25.05.1991

(51) INT CL[5]
H04B 3/20

(52) UK CL (Edition K)
H4R RDSD

(56) Documents cited
GB 2095519 A GB 2068698 A US 4864608 A

(58) Field of search
UK CL (Edition K) H4R RDSD RDSP
INT CL[5] H04B 3/20

(71) Applicant
Motorola Inc

(Incorporated in the USA - Delaware)

Corporate Offices, 1303 East Algonquin Road, Schaumburg, Illinois 60196, United States of America

(72) Inventor
Peter Kabell Jensen

(74) Agent and/or Address for Service
Hugh Christopher Dunlop
Motorola Limited, European Intellectual Property Operations, Jays Close, Viables Industrial Estate, Basingstoke, Hampshire, RG22 4PD, United Kingdom

(54) Enhancement of echo return loss

(57) A communications transceiver, such as a GSM mobile, is provided, for communicating in frames of encoded audio, comprising an audio input path (10), an audio output path (22), a voice activity detector (VAD 12) for detecting voice on the audio input path, and echo detecting means (25, 26, 27, 13, 28) for detecting unwanted echos on the input path resulting from acoustic coupling with the output path. Transmission of encoded audio from the audio input path is inhibited, by means (14, 29), in the presence of voice which is indicated as echo by the echo detecting means. The transceiver preferably has means for generating silence indicator (SID) frames in place of audio frames during periods of detected echo and means (40) Fig. 4 (not shown) for supplementing the silence indicator (SID) frames with comfort noise parameters and transmitting the supplemented frames for decoding as audio frames at a receiver.

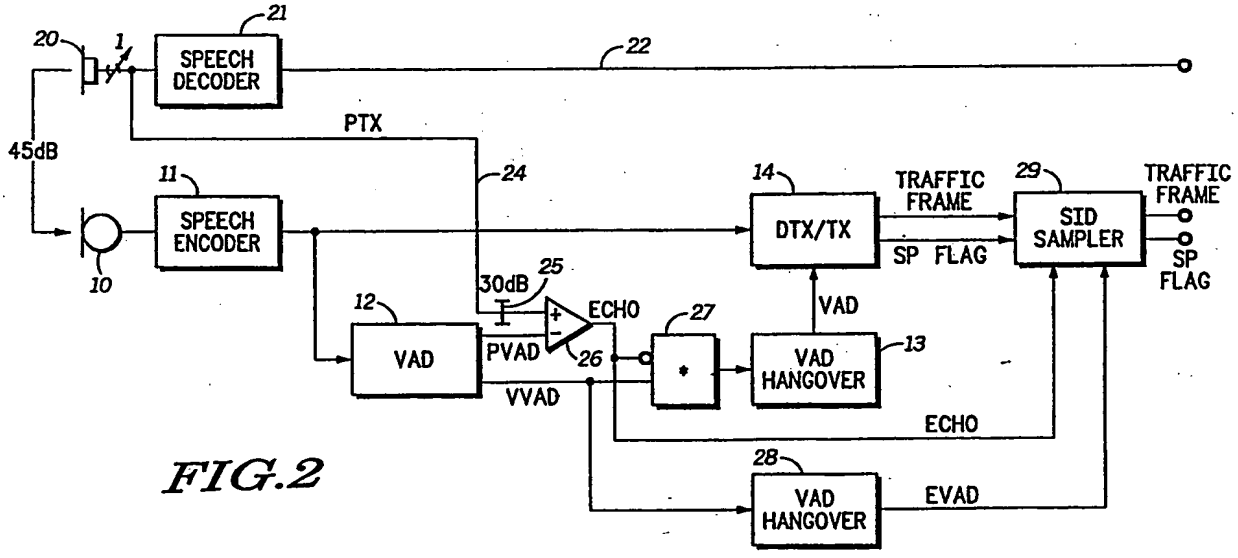

FIG.2

GB 2 256 351 A

At least one drawing originally filed was informal and the print reproduced here is taken from a later filed formal copy.

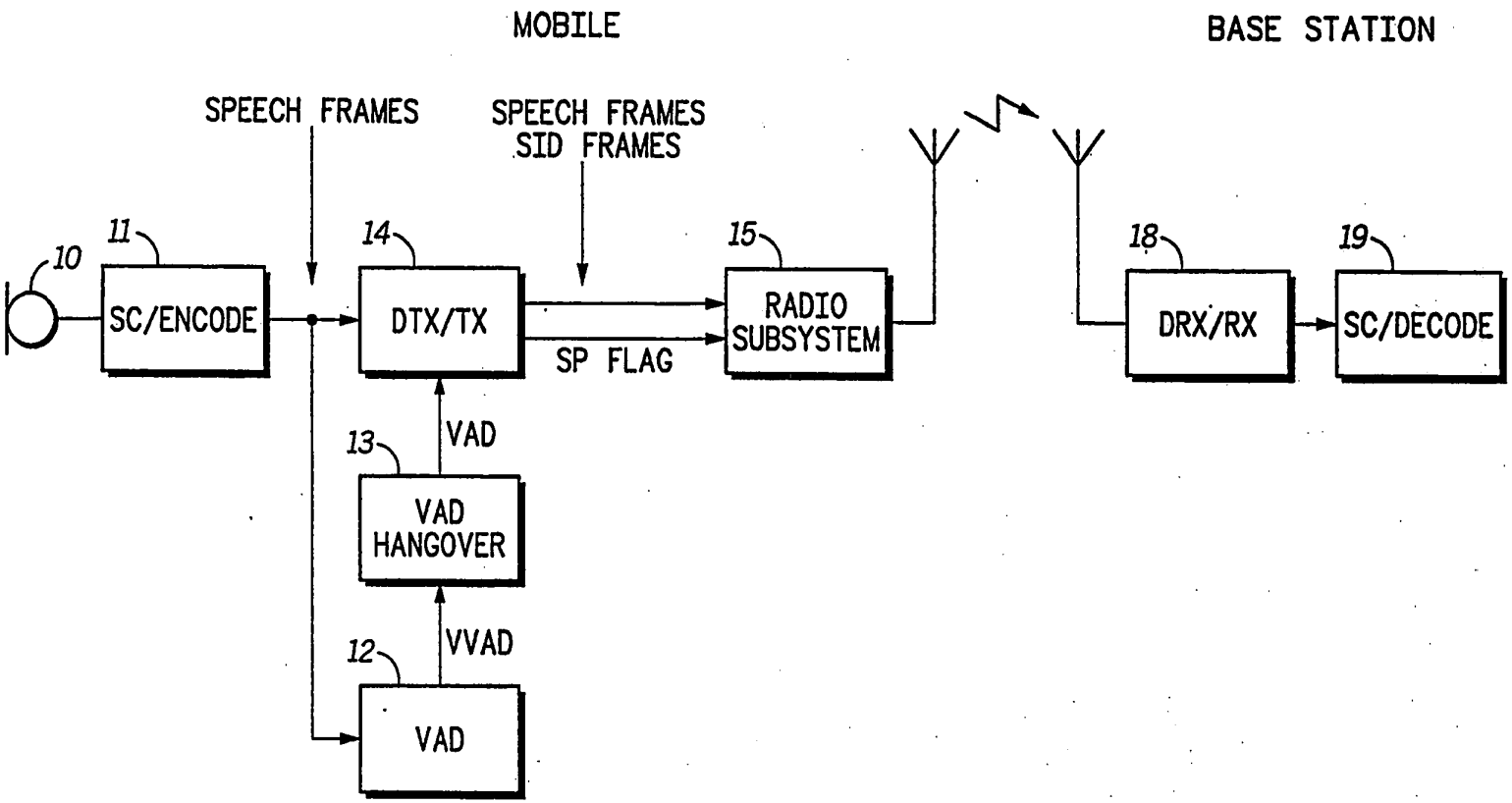

FIG.1
-PRIOR ART-

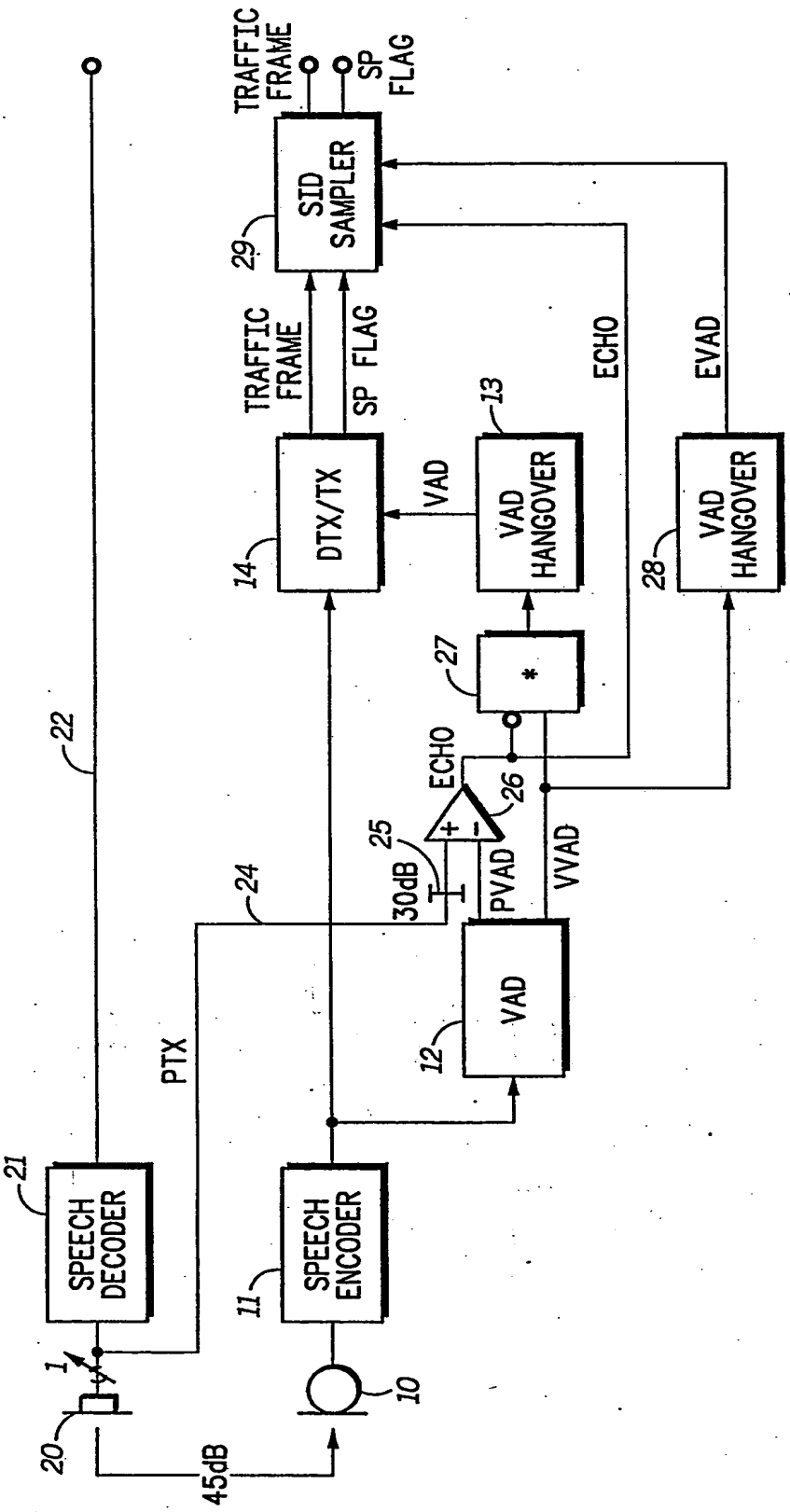
TRAFFIC FRAME
SP FLAG
SID SAMPLER
29
TRAFFIC FRAME
SP FLAG
ECHO
EVAD
13
VAD HANGOVER
VAD
DTX/TX
14
28
VAD HANGOVER
27
*
22
ECHO
26
25
30dB
PVAD
VVAD
24
VAD
12
PTX
SPEECH DECODER
21
SPEECH ENCODER
11
1
10
20
45dB

FIG.2

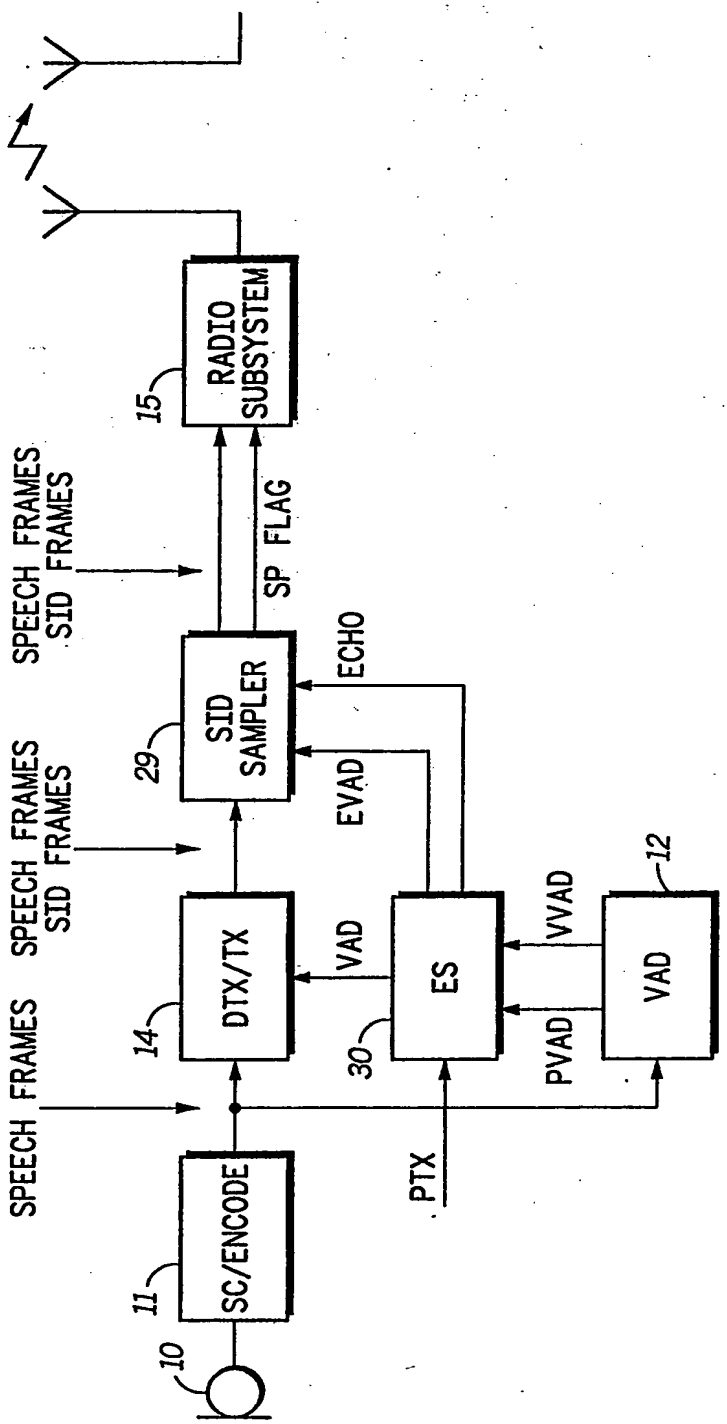

FIG.3

ECHO SUPPRESSOR FUNCTION WITH DTX

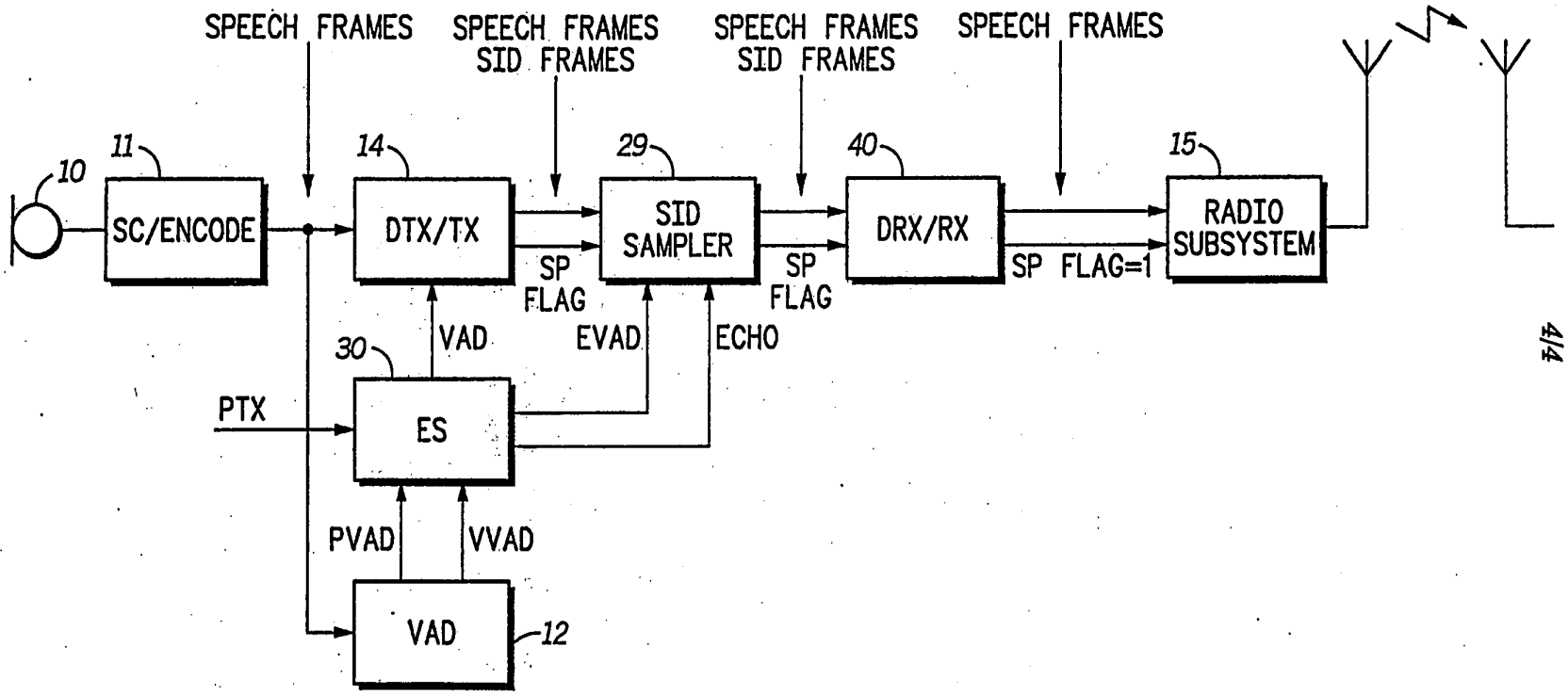

FIG.4

ECHO SUPPRESSOR FUNCTION WITHOUT DTX

ENHANCEMENT OF ECHO RETURN LOSS

Background of the Invention

This invention relates to a communication systems, such as a cellular radio system having a transmitter and a receiver and means for indicating the presence of voice in the signal transmitted from the transmitter to the receiver.

Summary of the Prior Art

In the GSM cellular radio system, speech is transmitted in frames of encoded speech data. The system defines that each mobile radio unit has a voice activity detector (VAD) for detecting voice on a channel. This is used for operation in a discontinuous transmission mode (DTx). When the DTx mode is inactive, frames of encoded speech are transmitted continuously, regardless of whether there is voice on the channel. When there is no voice on the channel, the speech coder encodes the background noise as if it is speech and this is transmitted and reproduced at the receiver end. When the DTx mode is active, and there is no voice on the channel, the absence of voice is detected by the VAD and, instead of transmitting the information required for full coding of each frame, the mobile transmits only the filter parameters for the voice decoder. A frame of filter parameters only is called a Silence Descriptor (SID) frame. The DTx operation is described in greater detail in GSM recommendation 06.31. The receiver fills in the rest of the frame with random data ("comfort noise") as described in GSM recommendation 06.12. Instead of transmitting every frame, the spectrum of the background noise is only updated every 24th frame. This reduces the overall level of radio activity in the system and therefore reduces co-channel interference. It also saves mobile battery life.

Throughout this description, the expression "DTx mode" will be used to refer to the mode in which SID frames are transmitted to the base station in the absence of voice.

This mode is entered on receipt of a command from the base station. The expression "DTx function" will be used to refer to the change at the mobile from voice operation to transmission of SID frames when in the DTx mode. The Base station recognizes the SID frames by looking at the excitation information. If this is set to zero the frame is regarded as a SID frame.

In the initial years of operation of the system, it is not intended that the DTx mode will be used. When overall traffic increases over time, the DTx mode will be activated. All mobiles, however, must have the ability to operate in both modes from the outset. In remote areas, it may not be necessary to activate the DTx mode. There will therefore be a situation where a mobile may have to switch modes depending on its location. This switching of modes is defined in the GSM specification.

Elements of a GSM system as described above are shown in Fig. 1. Referring to that figure, elements of a mobile radio telephone are shown on the left hand side and elements of a base station on the right hand side. The mobile comprises: a handset, of which a microphone 10 is shown; a speech coder 11 for encoding speech prior to transmission; a voice activity detector (VAD) 12 which detects the presence of voice and distinguishes voice from background noise; a VAD hangover element 13 for extending the VAD indicator; and a discontinuous transmitter DTx/tx 14 for generating SID frames when the VAD flag is cleared. The DTx/tx unit also provide a flag (SP flag) to the radio subsystem indication whether the frame is a speech frame or a SID frame. The speech SID frames are transmitted by means of a radio subsystem 15 to the base station. The base station has a discontinuous transmission receive (DTx/rx) element 18 and a speech decoder 19.

When in the DTx active mode, the VAD 12 detects the absence of voice on the channel and causes DTx/tx 14 to generate a SID frame. The SID frame contain only the speech coder filter parameters and zeros in place of excitation information. In this state, the mobile transmits silence

descriptor (SID) frames comprising only the speech coder filter parameters and zeros in place of excitation information. The base station receives the SID frames and replaces the excitation information with comfort noise from a comfort noise generator in the DTx/rx element 18. In this way, the comfort noise generator reconstructs background noise similar to that being received at the mobile microphone 10. Thus, the listener at the base station end does not have the disturbing affect of hearing voice and then suddenly hearing nothing.

A problem with mobile radios is acoustics feedback (echo return) from the mobile earpiece (not shown) to the microphone 10. The GSM cellular radio system has strict requirements for echo return loss. The specification requires that the echo return loss is less than or equal to 46 dB. With a good mechanical design of handset, the acoustic coupling from the earpiece to the microphone should be able to meet this requirement of a 46 dB difference. It is desirable, however, to provide a volume control to increase the handset audio output by 10 dB. When the earpiece output is increased by 10 dB, the echo return loss is only 36 dB. In addition, the maximum echo delay for an acoustic signal originating from a person connected to the public services telephone network (PSTN) and being returned as echo by virtue of acoustic coupling between the earpiece and the microphone should be 180 milliseconds. In practice, substantially greater delays will be introduced, depending on the transmission route. For example, a satellite connection could introduce delays of 400 milliseconds. This echo is disturbing to the ear of the receiver connected to the PSTN.

Summary of the Invention

According to the present invention, there is provided a communications transceiver, for communicating in frames of encoded audio, comprising an audio input path, an audio output path, a voice activity detector (VAD) for detecting voice on the audio input path, echo detecting means for

detecting unwanted echos on the input path resulting from acoustic coupling with the output path, and means responsive to the echo detecting means for inhibiting transmission of encoded audio from the audio input path in the presence of voice which is indicated as echo by the echo detecting means.

The transceiver preferably has means for generating SID frames in place of audio frames during periods of detected echo and means for supplementing the SID frames with comfort noise parameters and transmitting the supplemented frames for decoding as audio frames at a receiver. In this way, frames of echo are replaced with frames of comfort noise, even when there is no facility at the receiver for decoding of SID frames.

Brief Description of the Drawings

Fig. 1 shows elements of a mobile transceiver and a base station transceiver in accordance with the prior art.

Fig. 2 shows a communications transceiver in accordance with the preferred embodiment of the invention.

Fig. 3 illustrates the operation of a preferred embodiment when a DTx mode is active and

Fig. 4 shows the operation of a preferred embodiment of the invention when the DTx mode is not in use.

Description of the Preferred Embodiment

Referring again to the prior art arrangement of Fig. 1, more detailed explanation of the operation is as follows. When the VAD 12 detects speech (VAD = 1) coded speech is transmitted, and when no speech is detected (VAD = 0) the radio transmitter is switched off, except for short intervals when SID frames of comfort noise filter parameters are transmitted. SID frames are sent to the radio subsystem 15 as long as the VAD flag is low, but are only updated to the base station during a short interval following the speech burst and for every 24th frame. The DTx/rx software 18 in the base station receives the SID frames and inserts random

numbers in place of the missing excitation information, thereby generating comfort noise which resembles the noise at microphone 10. The comfort noise frames are decoded by the speech decoder 19.

The VAD function is based on an adaptive filter to increase the speech/noise ratio. The power of the filtered signal is compared with a threshold. Speech is indicated (VVAD = 1) whenever the threshold is exceeded. The VAD hangover element 13 is used to eliminate mid-burst and end-burst clipping of speech.

The VAD filter coefficients are obtained during speech pauses. In order not to update the filter coefficients during speech, the VAD is fairly sensitive, and an echo 45 dB under normal speech level can be detected in a quiet environment such as an office. This means that DTx/tx function 12 will open the up link radio channel and the echo will be transmitted. In a hand held-portable, a further problem arises from the transmission of the echo. As the radio path is opened during both near-and-end far-end talk, the DTx function will not reduce the power consumption of the mobile and the operation time will thus be reduced.

To remove the echo, it is not sufficient merely to adjust the VAD sensitivity such that an echo is not detected. Such an arrangement would introduce the problem that the VAD will update its filter coefficients during the echo and this will result in a reduction of the adaptive filter's ability to remove noise from the received signal. Moreover, if the comfort noise parameters (which are updated every 24th frame) were to be updated while an echo is present, this would result in very fluctuating comfort noise, which would be disturbing to the listener at the base station end.

Referring to Fig. 2, there is shown, in addition to the elements of Fig. 1, an earpiece 20 and speech decoder 21 on an audio output path 22. Furthermore, the following new elements are shown: an audio output path feedback line 24, an attenuator 25, a comparator 26, a logic gate 27, a further VAD hangover element 28 (in addition to the VAD hangover element 13) and a silence descriptor sampler 29.

The echo suppressor function in accordance with the present invention may be used with or without the DTx function. For the purposes of the present explanation, it will be assumed that the DTx function is active.

The power of the received signal Ptx on feedback line 24 and the near-end signal Pvad from the VAD 12 are compared (as already described, the near-end signal power is already calculated on the basis of a noise filtered signal). The comparison is made in comparator 26 after attenuation of the PTx signal in the attenuator 25. The ECHO flag is set accordingly to:

If Ptx-30dB > Pvad; ECHO=1 (Echo detected)

If Ptx-30dB < Pvad; ECHO=0

Since the echo loss in a normal handset is about 45 dB the figure of 30dB allows a good margin for deviations in the send and receive loudness ratings. If a volume control 23 is present, the threshold of 30dB can be adjusted according to the volume of setting. If the volume is increased the threshold should be decreased.

The VAD flag indicating whether to send coded speech frames is modified according to the following:

VAD = (VVAD*^ECHO)++

Where:

* = AND

^ = logical not

++ = Hangover period.

As can be seen from the table below. The VAD flag is only set when near-end or double talk is present

Near-end speech	Far-end speech	VVAD	ECHO	VVAD*^ECHO
0	0	0	0	0
0	1	Ø	1	0
1	0	1	0	1
1	1	1	0	1

If only far-end speech is present the VVAD may and may not detect the echo depending on noise and volume settings, this is indicated by the don't care state Ø.

The EVAD flag is set according to:

EVAD = (VVAD++). This flag is therefore identical to the original VAD flag of the prior art. The SID sampling unit 29 reads the ECHO and the EVAD flags and decides when to update the comfort noise parameters.

The comfort noise parameters are only sampled just after a speech burst and for every 24th frame. Between updates of the SID frames the old sample is repeatedly sent to the radio subsystem. This updating is equal to the sampling performed by the radio subsystem, however the update of the SID frames for every 24th frame are not simultaneous.

The comfort noise parameters are only updated whenever eVAD = 0 and ECHO = 0. This avoids updating of the parameters during an echo and therefore avoids unnecessary perturbation of the comfort noise at the receive end. To take into account situations where very loud far-end noise may cause the ECHO bit to be set for long periods or situations where no pauses exist between far-and and near-end speech an additional strategy is used to ensure the best quality of comfort noise.

If eVAD = 1 at the sampling time, this indicates that an echo was detected by the VAD and the comfort noise parameters are not updated. If eVAD = 0 and ECHO - 1 an echo was detected by the echo suppressor but not by the VAD. The ECHO may be very low speech or very loud background noise. The comfort noise parameters are not updated unless this situation is repeated three times in a row not interrupted by any other combination of the two flags eVAD and ECHO. This indicates that the echo was background noise and the comfort noise parameters are updated hereafter as long as eVAD = 0 and ECHO = 1.

Furthermore, the comfort noise parameters are only updated just after a speech burst and for every 24th frame. Between updates of the SID frames, the old sample is repeatedly sent to the radio subsystem. This updating is

equal to the sampling performed by the radio subsystem, however the updates of the SID frames for every 24th frame are not simultaneous.

The above described system with the DTx mode active is shown in block diagram form in Fig. 3. The elements making up the echo suppressor (attenuator 25, comparator 26, logical gate 27 and hangover elements 13 and 28) are all shown as a single block 30. Referring to the transmit path from the microphone horizontally through the diagram to the radio subsystem 15, it is shown that speech frames emerge from the speech coder 11 and speech frames interspersed with SID frames emerge from the DTx/tx 14 and the SID sampler 29.

Referring now to Fig. 4, the arrangement of Fig. 3 is shown with a modification in accordance with the invention for use when the DTx mode is inactive (for example in the initial years of operation the GSM system, or in remote areas).

Because the DTx mode is inactive, the DTx/rx element 18 has been removed from the part of the diagram illustrating the base station. In practice, this element may be present but merely be inactive. In accordance with the invention, a further element has been included in the mobile in the form of DTx/rx element 40. This element is simple to implement, because it merely requires re-use of the DTx/rx software in the mobile. DTx is allowed both in the uplink direction (mobile to base station) and the downlink (Base station to mobile). Element 40 is activated when the DTx mode is inactive, and de-activated when a DTx activate signal is received from the base station.

The base station informs the mobile if DTx is allowed to uplink direction. The base station may use DTx in the downlink direction, the DTx/rx unit in the mobile is therefore always activated.

The DTx/rx element 40 receives SID frames from the SID sampler 29 and inserts random numbers in place of the missing excitation information. In this way, frames of comfort noise are generated which can be treated as speech frames. For the purposes of operation of the system, therefore, the frames

emerging from DTx/rx 40 are exclusively speech frames and these are transmitted by the radio subsystem 15 to the base station for decoding.

For the purposes of implementation, time alignment flags required for the DTx/rx unit 40 are supplied by the SID sampler. The DTx/rx unit 40 handles interpolation of the SID frames.

When the echo suppressor is working without the DTx mode, it prevents the DTx/tx unit from opening the radio path if an echo is mistaken for speech by the VAD unit. This eliminates the echo and increases operation time in a hand portable.

The echo suppressor prevents the DTx/tx function from updating the comfort noise parameters when an echo is present. This advantage is derived in both operational modes of the system. The quality of the comfort noise is thereby increased.

The arrangement described cancels echos in situations where an echo canceller *per se* is unsuitable, because it would require a substantial amount of computation time and because of the low level of echo energy.

If the mobile is operating in hands-free mode, voice switching with comfort noise insertion as described can be used in conjunction with an echo canceller to ensure sufficient echo loss.

GLOSSARY OF TERMS

ABBREVIATIONS

DTx/rx	Discontinuous transmission receive unit. If a speech frame is received the DTx unit sends this frame to the speech decoder. If a SID frame is received the zeros in the frame is filled with random numbers. Between updates of the SID frame the old SID frame is used. Between updates of the SID frame no information is received by the DTx/rx unit.
DTx/tx	Discontinuous transmission transmit unit. This unit receives a coded speech frame from the speech coder and a VAD flag from the VAD unit. If the VAD flag is set, coded speech is sent. If the VAD flag is cleared, the excitation information is set to zero. The resulting frame is called a SID frame. The SID frames are only transmitted every 24th frame. The rest of the time the transmitter is shut off.
ECHO	ECHO flag. Flag indicating whether only far end speech is present. This tells the system not to transmit speech frames, as the frames may contain an echo. The flag is set by comparing Pvad and Ptx.
eVAD	Copy of the VVAD flag before it is modified by the echo suppressor. This flag tells the echo suppressor that the comfort noise should not be updated, as the frame contains speech or an echo. A hangover period is added to the VVAD flag.

Frame	Time interval of 20 msec (160 samples of 13 bit uniform coded samples - 1280 bits) corresponding to the time segments of the speech transcoder, also used as a short term for a traffic frame.
Ptx	Power of far end signal. This signal is not noise filtered.
Pvad	Power of near end signal. The near end signal is noise filtered and the power is measured.
SID	Silence Descriptor frame. Frame containing only the spectrum information. The excitation of the filter is set to zero. The DTx/rx unit recognises the SID frame and fills in random numbers instead of the zeros. This is equivalent to exciting the synthesis filter with a random generator. The modified SID frame is repeatedly sent to the speech decoder 24 times, until a new SID frame is received.
SPflag	Internal flag in the mobile indication to the radio subsystem whether the traffic frame is speech-frame or a SID-frame.
TAF	Time Alignment Flag. If the TAF flag is set the DTx/rx unit looks for SID frames, SID frames are only updated at certain times known to mobile and the base station.
Traffic Frame	Block of coded speech. 260 information bits. In principal the speech is coded as a synthesis filter (spectrum) and an excitation of the filter.

VAD	The modified VVAD flag with a hangover period at the end of the speech burst.
VVAD	Flag indicating whether speech is present in the frame. This flag may be cleared by the echo suppressor if the speech in the frame is an echo.

CLAIMS

1. A communications transceiver, for communicating in frames of encoded audio, comprising an audio input path (10), an audio output path (22), a voice activity detector (VAD) for detecting voice on the audio input path, echo detecting means (30) for detecting unwanted echos on the input path resulting from acoustic coupling with the output path and means (14. 29) responsive to the echo detecting means for inhibiting transmission of encoded audio from the audio input path in the presence of voice which is indicated as echo by the echo detecting means.

2. A communications transceiver according to claim 1, further comprising comfort noise generating means (40) for transmitting comfort noise parameters in place of encoded audio from the audio input path, for decoding as normal speech at a receiver.

3. A communications transceiver according to claim 2, further comprising DTx mode activation means for receiving and decoding a DTx mode authorization signal, wherein the comfort noise generating means (40) is responsive to the DTx mode activation means and activated in the absence of such an authorization signal.

4. A communications transceiver according to claim 1, operable in a discontinuous transmission (DTx) inactive mode, in which substantially all frames of audio transmitted are complete frames of encoded audio, and operable in a DTx active mode in which silence indicator descriptor frames are transmitted for reproduction of periods of absence of speech, further comprising:

means (14), operable in the DTx inactive mode, for generating silence indicator descriptor parameters in response to the detection of echo by the echo detector,

means (40) for supplementing the SID parameters with comfort noise parameters to generate complete frames of encoded audio,

means (15) for transmitting the complete frames of encoded audio and

means (29) for repeating the silence indicator parameters from frame-to-frame during the presence of echo and the absence of voice.

5. A communications transceiver according to any one of the preceding claims, wherein the echo detecting means comprise comparator means (25, 26) for comparing the signal on the audio input path with the signal on the audio output path (22) when attenuated by a selectable attenuation; wherein volume control means (23) are provided for controlling the volume on the output path; and wherein the selectable attenuation of the output used for comparison is a dependent on the setting of the volume control.

6. A communications transceiver according to claim 2 or 3, further comprising means for generating comfort noise parameters and updating them at update periods spanning a number of frames, further comprising means for suppressing the updating of the comfort noise parameters for a predetermined number of update periods while a potential echo is detected by the echo detecting means and no voice is detected by the VAD, and updating the parameters after said predetermined number of frames.

15.

. atents Act 1977
Examiner's report to the Comptroller under Section 17 (The Search Report)

Application number
9111380.3

Relevant Technical fields

(i) UK Cl (Edition K) H4R (RDSD,RDSP)

(ii) Int Cl (Edition 5) H04B 3/20

Search Examiner
K WILLIAMS

Databases (see over)

(i) UK Patent Office

(ii)

Date of Search
22 OCTOBER 1991

Documents considered relevant following a search in respect of claims 1-6

Category (see over)	Identity of document and relevant passages	Relevant to claim(s)
A	GB 2095519 A (WESTERN ELECTRIC) - see Figure 1	1
A	GB 2068698 A (WESTERN ELECTRIC) - see Figure 6	1
A	US 4864608 A (HITACHI) - see column 5, lines 5-13	1

Category	Identity of document and relevant passages	Relevant to claim(s)

Categories of documents

X: Document indicating lack of novelty or of inventive step.

Y: Document indicating lack of inventive step if combined with one or more other documents of the same category.

A: Document indicating technological background and/or state of the art.

P: Document published on or after the declared priority date but before the filing date of the present application.

E: Patent document published on or after, but with priority date earlier than, the filing date of the present application.

&: Member of the same patent family, corresponding document.

Databases: The UK Patent Office database comprises classified collections of GB, EP, WO and US patent specifications as outlined periodically in the Official Journal (Patents). The on-line databases considered for search are also listed periodically in the Official Journal (Patents).

Published 1992 at The Patent Office, Concept House, Cardiff Road, Newport, Gwent NP9 1RH. Further copies may be obtained from Sales Branch, Unit 6, Nine Mile Point, Cwmfelinfach, Cross Keys, Newport, NP1 7HZ. Printed by Multiplex techniques ltd, St Mary Cray, Kent.